## AMENDMENTS TO THE CLAIMS

**This listing of claims will replace all prior versions and listings of claims in the application:**

## LISTING OF CLAIMS:

1.-19. **(Canceled).**

20. **(New)** A method for identifying a candidate protein useful as an anti-infective, comprising:

(a) calculating computationally protein sequence-based attributes from all protein sequences of a pathogenic organism, wherein said protein sequence-based attributes are selected from a group consisting of percentage of charged amino acids, percentage hydrophobicity, distance of protein sequence from a fixed reference frame, measure of dipeptide complexity, and measure of hydrophobicity from a fixed reference frame;

(b) clustering computationally said all protein sequences based on said protein sequence-based attributes using Principle Component Analysis;

(c) identifying computationally outlier proteins, wherein said outlier proteins appear outside a main cluster;

(d) comparing said outlier proteins to known proteins to identify a unique outlier protein; and

(e) validating said unique outlier protein as an anti-infective.

21. **(New)** The method of claim 20, wherein said pathogenic organism is selected from the group consisting of *B.burgdorfei*, *C.jejuni*, *C.pneumoniae*, *C.trachomatis*, *H.influenzae*, *H.pylori*, *L.major*, *M.genitalium*, *M.pneumoniae*, *M.tuberculosis*, *N.menigitis*, *P.aeruginosa*, *P.falciparum*, *R.prowazekii*, *T.pallidum*, and *V.cholerae*.

22. **(New)** The method of claim 20, wherein said protein sequence-based attributes are selected from the group consisting of fixed protein attributes and variable protein attributes.

23. **(New)** The method of claim 22, wherein a variable protein attribute is a distance of protein sequence from a variable reference frame.

24. **(New)** The method of claim 20, wherein said clustering is done by Principle Component Analysis using correlation coefficient between said protein sequence-based attributes.

25. **(New)** The method of claim 20, wherein said clustering is based upon analysis of protein sequence-based attributes and not based upon sequence pattern linked to biochemical functions.

26. **(New)** The method of claim 20, wherein said unique outlier protein is non-homologous to known anti-infective proteins from a pathogen selected from the group consisting of *B.burgdorfei*, *C.jejuni*, *C.pneumoniae*, *C.trachomatis*, *H.influenzae*, *H.pylori*, *L.major*, *M.genitalium*, *M.pneumoniae*, *M.tuberculosis*, *N.menigitis*, *P.aeruginosa*, *P.falciparum*, *R.prowazekii*, *T.pallidum*, and *V.cholerae*.

27. **(New)** The method of claim 20, wherein said unique outlier protein has an amino acid sequence selected from the group consisting of SEQ ID Nos: 1-31.

28. **(New)** The method of claim 20, wherein said unique outlier protein has an amino acid sequence selected from the group consisting of SEQ ID Nos: 32-118.

29. **(New)** The method of claim 20, wherein steps are performed by a computer system comprising:

(1) a central processing unit (CPU), wherein said CPU executes DISTANCE program and clusters protein sequences based on protein sequence-based attributes using Principle Component Analysis, thereby producing results;

(2) a memory device accessed by said CPU, wherein said memory device stores said results;

(3) a display on which said CPU displays said results in response to user inputs; and

(4) a user interface device.

30. **(New)** The method of claim 20, wherein said unique outlier protein may be used for a diagnostic purpose.

31. **(New)** The method of claim 20, wherein said unique outlier protein may be used as a vaccine candidate.

32. **(New)** The method of claim 20, wherein said unique outlier protein may be used for a therapeutic purpose.

33. **(New)** The method as of claim 20, wherein said unique outlier protein can elicit an immune response.